مجلۀ زبانشناسی و گویش‌های خراسان ، سال دوازدهم ، شماره ٤، پاییز زمستان ١٣٩٩، شمارۀ پیاپی ٢٣

doi: 10.22067/jlkd.2021.68589.1012

تحلیل عبارات احتیاط‌آمیز در پیکره زبانی آکادمیک فردوسی: مقایسه‌ی دو حوزه علوم‌انسانی و علوم‌پایه

عطیه کامیابی‌گل (استادیار گروه آموزشی زبان و ادبیات فارسی و گروه زبان‌شناسی، دانشگاه فردوسی مشهد، مشهد، ایران) (نویسنده مسئول)

الهام اخلاقی باقوجری (دکترای زبان‌شناسی همگانی، دانشگاه فردوسی مشهد، مشهد، ایران)

هانیه حبیبی (دانشجوی دکتری تکنولوژی ترجمه زبانشناسی، دانشگاه ژنو،ژنو، سوئیس)

صص: ٢٤٦-٢٢٧

چکیده

پیکره شامل متن‌های پیوسته طبیعی است که به صورت الکترونیکی ارائه شده‌است و می‌توان از آن اطلاعات درباره عناصر زبانی، به‌صورت واژگانی و غیرواژگانی در کم‌ترین زمان ممکن و با بالاترین میزان دقت، استخراج نمود. نویسندگان این مقاله، پیکره زبانی آکادمیک فردوسی را ایجاد کرده و هدف از این مقاله معرفی پیکره اعضای هیات علمی دانشگاه فردوسی مشهد و به‌کارگیری آن در تحلیل و توصیف عبارات احتیاط‌آمیز در دو حوزه علوم‌انسانی و علوم‌پایه است. عبارات احتیاط‌آمیز میزان قطعیت کلام نویسنده را با کلماتی مانند «شاید»، «ممکن است» و «حتما» مشخص می‌کند. پژوهش‌هایی که تاکنون بر روی این عبارات انجام شده‌است، با داده‌هایی اندک و به صورت کاملا دستی انجام‌شده‌اند. در این پژوهش ضمن استفاده از دسته‌بندی سلاجرمیر (١٩٩٧) برای شناسایی عبارات احتیاط‌آمیز، داده‌ها با استفاده از پیکره الکترونیکی ساخته‌شده که شامل ١١٠٠ مقاله زبان فارسی می‌باشد، استخراج شدند. بر اساس نتایج حاصل از این پژوهش میزان کاربرد عبارات احتیاط‌آمیز در حوزه علوم‌انسانی حدود دو برابر بیشتر از حوزه علوم‌پایه بود. در این میان استفاده از شرطی‌ها و عبارات حاکی از شک و تردید در هر دو حوزه بسامد

تاریخ دریافت: ١٣٩٩/١١/٠٩ تاریخ پذیرش: ١۴٠٠/٠٢/٢١

پست الکترونیکی:
١. kamyabigol@um.ac.ir
٢. Elhamakhlaghi80@gmail.com
3. hanieh.habibi@unige.ch

بالایی داشتند. نتایج حاصل از این پژوهش به علت بالاتر بودن حجم داده‌های مورد استفاده نسبت به تحقیقات پیشین، قابلیت تعمیم بیشتری به عنوان الگو در نوشتن مقالات علمی زبان فارسی را دارا هستند. همچنین تأیید شد پیکره ساخته‌شده در این پژوهش که تا ۹٦ درصد دقت دارد، به عنوان منبع بسیار خوبی برای تحلیل و تحقیق بر روی متون علمی قابل استفاده است.

واژه‌های کلیدی: پیکره، مقالات علمی- پژوهشی، زبان فارسی، اعضای هیات علمی، عبارات احتیاط‌آمیز

۱-مقدمه

بر اساس تعریف لیکاف[۱](۱۹۷۲)، «عبارات احتیاط‌آمیز[۲]کلماتی هستند که کارشان مبهم کردن موضوع است» (لیکاف، ۱۹۷۲:۱۹۵). عبارات احتیاط‌آمیز انواعی دارد و پژوهش‌گران مختلف دسته‌بندی‌های متفاوتی برای آن معرفی کرده‌اند. از جمله کلمن[۳](۱۹۷۷) که تقسیم-بندی خود را به این صورت در نظر گرفته است: عبارات احتیاط‌آمیزِ معمول، مرکب، حروف ربط تباینی، بند شرط، ادات، ساخت مجهول، قضاوت شخصی و قضاوت درباره ارزش صدق. هایلند[٤](۱۹۹۵) برای عبارات احتیاط‌آمیز، دسته‌بندی هشت‌مقوله‌ای دارد، از جمله افعال واژگانی (نشان می‌دهد)، ساختارهای قیدی (احتمالاً)، افعال کمکی (ممکن است)، صفات کمکی (بیشترین)، اسامی کمکی (تخمین)، اقرار به نداشتن دانش کافی (نمی‌دونم که آیا..)، اشاره به محدودیت شرایط (اگر این انگاره درست باشد) و اشاره به مدل، نظریه و انگاره (نتیجه‌گیری مشکل است) (هایلند، ۱۹۹۵، ص. ۳٦). سلاجرمیر[٥](۱۹۹۷) نیز تقسیم‌بندی دارد که مشتمل بر هفت دسته عمده است که در این پژوهش از همین دسته‌بندی استفاده شده‌است.

---

۱ Lakoff
۲ Hedges
۳ Clemen
٤ Hyland
٥ Salager Meyer

بعلاوه جلیلی‌فر[۶] (۲۰۰۶) موارد زیر را به عنوان انگاره عبارات احتیاط‌آمیز معرفی کرده است: افعال معین، افعال وجهی، گروه‌های اسمی، قیدی، صفتی، وجهی، تقریبی‌های درجه، کمیت، بسامد و زمان، جملات شرطی، عبارات احتیاط‌آمیز مرکب، مجهول بدون ذکر کننده کار و افزودن عناصر خاص به صفت.

باید خاطر نشان کرد استفاده از این نشانه‌ها نقش عمده‌ای در گفتمان علمی داشته و در واقع ابزاری است که بیانگر تردید نویسنده و عدم قطعیت گزاره بوده و نویسنده با استفاده از آن‌ها می‌تواند عدم پایبندی کامل به ارزش صدق یک گزاره را نشان دهد. از طرفی امروزه نیاز به افزودن اطلاعات جانبی به داده‌های متنی عظیم جهت پردازش آن‌ها توسط ماشین بیش از پیش احساس می‌شود. انجام این کار با ساختن پیکره و برچسب‌گذاری (حاشیه‌نویسی[۷]) آن که در قالب‌های مختلف می‌تواند صورت بگیرد، باعث غنای داده‌های موجود و آماده شدن آن‌ها برای استفاده در تحقیقات و جستجوهای علمی می‌شود. این نشانه‌ها در زبان‌های دیگر به ویژه زبان انگلیسی از اهمیت بالایی برخوردار هستند و در پیکره‌های مختلفی بررسی شده‌اند، اما در زبان فارسی تحقیق‌های انجام شده بر روی مجموعهٔ محدودی از داده‌ها و به صورت دستی انجام شده‌است. در این پژوهش برای بررسی عبارات احتیاط‌آمیز از پیکره تهیه شده توسط نویسندگان با مجموع ۷۱۵٤۲۰۲ واژه که مشتمل بر ۱۱۰۰ مقاله از پژوهش‌های اعضای هیات علمی دانشگاه فردوسی مشهد است، استفاده شد و با توجه به تقسیم‌بندی آموزشی دانشگاه و امکاناتی که در استفاده از پیکره وجود داشت دو حوزه آموزشی متفاوت (علوم‌انسانی و علوم‌پایه) برای بررسی و مقایسه عبارات احتیاط‌آمیز انتخاب گردید.

سوال‌هایی که در این پژوهش به دنبال پاسخ مناسبی برایشان هستیم بدین ترتیب هستند:

۱. نوع عبارات احتیاط‌آمیز بکار رفته در مقالات علمی-پژوهشی اعضای هیات علمی حوزه علوم‌انسانی چگونه است؟

۲. میزان استفاده عبارات احتیاط‌آمیز بکار رفته در مقالات علمی-پژوهشی اعضای هیات علمی حوزه علوم‌انسانی چگونه است؟

---

۶. Jalilifar

۷. Annotation

۳. نوع عبارات احتیاط‌آمیز بکار رفته در مقالات علمی-پژوهشی اعضای هیات علمی حوزه علوم‌پایه چگونه است؟

٤. میزان استفاده عبارات احتیاط‌آمیز بکار رفته در مقالات علمی-پژوهشی اعضای هیات علمی حوزه علوم‌پایه چگونه است؟

مبانی نظری

پژوهش حاضر بر اساس چارچوب نظری سلاجرمیر[۸](۱۹۹۷) و تقسیم‌بندی او از عبارات احتیاط‌آمیز، که مشتمل بر هفت دسته عمده است، انجام پذیرفت. در ادامه به این تقسیم‌بندی پرداخته می‌شود:

۱. **افعال کمکی**. مانند: خواستن یا توانستن.

۲. **افعال کنش‌گفتاری** که نشان‌دهنده شک و ارزیابی نویسنده است؛ مانند: نشان دادن، فرض کردن، معتقد بودن، پیشنهاد دادن.

۳. **عبارات اسمی، قیدی و صفتی** که خود شامل زیربخش‌های زیر می‌شود:

- **صفت‌های احتمال**. مانند: «ممکن» و «محتمل».
- *اسامی احتمال*. مانند: «فرض» و «احتمال».
- *قیدهای احتمال*. مانند: «به‌طور ممکن»، «به‌طور احتمالی» و «به‌طور فرضی».

٤. **تقریب‌های[۹]اندازه، زمان، کیفیت و بسامد**. مانند: «معمولا»، «به‌سختی»، «اغلب» و یا «حدودا».

۵. **عبارات آغازین جمله**. مانند: «بر اساس دانش ما»، «این نگرش ما است که» و یا «ما احساس می‌کنیم که».

٦. **جملات شرطی**.

۷. **عبارات احتیاط‌آمیز مرکب**. مانند: «ممکن است پیشنهاد شود»، «شاید منطقی به‌نظر برسد که این‌گونه فرض کنیم» و یا «ممکن است غیر محتمل به‌نظر برسد».

---

۸ Salager Meyer

۹ Approximators

مهم‌ترین دلیل انتخاب این دسته‌بندی کامل و جامع بودن آن است، به‌علاوه بیشترِ تقسیم‌بندی‌های بعد از آن به نوعی با این دسته‌بندی هم‌پوشانی دارند یا به‌طور مستقیم از آن اقتباس شده‌اند.

۲-پیشنه پژوهش

در ساخت پیکره مسایل مختلفی مطرح می‌شود، از جمله دسته‌بندی مقولات گفتار، ارزیابی برچسب‌گذاری متون زبان فارسی و سایر مقوله‌های زبانی؛ در حالی‌که برخی به ارزیابی برچسب‌گذاری متون زبان فارسی پرداخته‌اند. به عنوان نمونه فرخ (۱۳۸۱) دسته‌بندی مبسوطی از افعال زبان فارسی برای تشخیص خودکار رایانه ارائه کرده است و دانشکار (۱۳۸۱) نرم‌افزار تشخیص فعل را طراحی کرده است. اما از آن‌جا که تمرکز این پژوهش بر روی بررسی عبارات احتیاط‌آمیز است، در ادامه به بیان برخی از آثار موجود در این حوزه اشاره می‌شود.

بر اساس تعریف هایلند (۱۹۹۶b) عبارات احتیاط‌آمیز عبارتند از هر روش زبانی که برای اشاره به موارد زیر استفاده شود:

- اطمینان کامل نداشتن به صدق گزاره بیان شده،
- تمایل به بیان عبارات به صورت مقوله‌بندی‌نشده.

به‌علاوه این عبارات می‌توانند گزاره‌ای را از زبان نویسنده به عنوان عقیده به جای بیان واقعیت بیان کنند (هایلند، ۱۹۹۶b: ٤٧٨).

در این راستا نقش مهم استفاده از عبارات احتیاط‌آمیز در متون علمی و مقالات پژوهشی توسط مطالعات مختلف به‌خوبی مشخص شده است (سلاجرمیر، ۱۹۹٤؛ هایلند، ۱۹۹٦a، ۱۹۹۸a، ۱۹۹۸b، ۲۰۰۰؛ ورتالا[10]، ۲۰۰۱). همچنین پژوهشگران مختلفی علاقه‌مند به بررسی عبارات احتیاط‌آمیز در متون علمی در بین زبان‌های مختلف از جمله انگلیسی و فارسی بوده‌اند (هایلند و تسو[11]، ۲۰۰٤؛ مارتین و برگز[12]، ۲۰۰٤؛ ولد[13]، ۲۰۰٦؛ کنگ[14]، ۲۰۰٦؛ جلیلی‌فر، ۲۰۰۷، میلان[15]، ۲۰۰۸؛ سمایی، خسرویان و بغایری، ۲۰۱٤).

---

[10] Varttala

[11] Hyland & Tse

[12] Marti´n. & Burgess

[13] Wold

اما از میان پژوهش‌های انجام‌شده بر روی مقالات زبان فارسی، عبـدی (۲۰۰۲) بـا بررسـی ۵۵ مقاله تفاوت‌های چشم‌گیری در به‌کارگیری این ابزار در بین مقـالات چـاپ شـده دو رشـته علوم طبیعی و علوم اجتمـاعی مشـخص کـرده‌اسـت. فلاحتـی[6](۲۰۰٤) در پایـان‌نامـه خـود بـا بررسی ۲٤ مقاله به مقایسه عبارات احتیاط‌آمیز در متون علمی فارسی و انگلیسـی پرداخته‌اسـت و به این نتیجه رسیده است که بـه طـور کلـی در مقـالات انگلیسـی تعـداد کلمـات و عبـارات احتیاط‌آمیز بیشتر از زبان فارسی است و همچنین رشته روان‌شناسی در متـون انگلیسـی زبـان و رشته داروسازی در متون فارسی‌زبان بالاترین بسامد کاربرد این عبارات را دارند.

فریزر[7](۲۰۱۰) عبارات احتیاط‌آمیز را به عنوان راهکاری بلاغی به‌شـمار مـی‌آورد کـه بـرای کاهش اجبار در کنش‌های گفتاری بکار می‌رود. وی بیان می‌کند از عبارات احتیاط‌آمیز مـی‌تـوان در ارائه مقدمه و ترغیب خوانندگان، توصیف روش‌ها، بحث یافتـه‌هـا و نتـایج پـژوهش یـاری جست. تـاکی و رضـایی[8](۲۰۱٤) بـه مقایسـهٔ کـاربرد عبـارات احتیـاط‌آمیز در پایـان‌نامـه‌هـای انگلیسی و فارسی در مقاطع کارشناسی ارشد و دکتری ایرانیان پرداخته‌است. آن‌ها کـاربرد انـواع عبارات احتیاط‌آمیز را متفاوت ذکر کرده‌اند و بیان کرده‌اند بسامد کاربرد این عبـارات در پایـان-نامه‌های نوشته شده به زبان انگلیسی بالاتر بوده است.

صداقت، بی‌ریا و اسدی‌امیرآبادی[9](۲۰۱۵) به بررسـی بـین‌فرهنگـی عبـارات احتیـاط‌آمیز در ۳۰ سرمقاله فارسی و انگلیسی پرداخته‌انـد و بسـامد کـاربرد ایـن عبـارات را در سـرمقاله‌هـای انگلیسی بالاتر برآورد کرده‌اند. همچنین اعـلام کـرده‌انـد از بـین سـه نـوع سـرمقاله اجتمـاعی، سیاسی و اقتصادی در زبان انگلیسی، سـرمقاله‌هـای سیاسـی و در زبـان فارسـی، سـرمقاله‌هـای اقتصادی بالاترین حجم کاربرد عبارات احتیاط‌آمیز را داشته‌اند.

دوستی و اسلام‌راسخ[10](۲۰۱٦) به بررسی تفاوت‌های جنسـیتی دانشـجویان زبـان در کـاربرد عبارات احتیاط‌آمیز در تعاملات بین‌فردی پرداخته‌اند. آن‌ها بـه ایـن نتیجـه رسـیده‌اند کـه زنـان

---

4. Kong
5. Millan
6. Falahati
7. Fraser
8. Taki & Rezaei
9. Sedaghat, Biriya & Asadi Amirabadi
10. Dousti & Eslami Rasekh

بیشتر از مردان از این‌گونه عبارات استفاده می‌کنند. همچنین زنـان بـرای دوسـتانه بـودن روابـط تمایل به پذیرش نظرات دیگران دارند، اما آن‌ها کـاربرد زیـاد عبـارات احتیـاطآمیز را حـاکی از پایین‌تر بودن جایگاه اجتماعی خود نمی‌دانند.

هاشمی و شـیرزادی[۱](۲۰۱٦) بـا بررسـی عبـارات احتیـاطآمیز در ۱۵۰ مقالـه حـوزه زبـان-شناسی کاربردی با سه روش تحقیق متفاوت کمّی، کیفی و ترکیبی، این‌گونه نتیجه گرفته‌انـد کـه در این عبارات در بخش بحث و بررسی مقالات کمّـی بـالاترین بسـامد را دارا هسـتند. سـپس مقالات با روش ترکیبی و در انتهای پیوستار مقالات با رویکرد کیفی قـرار دارنـد. همچنـین بـر اساس معیـار هایلنـد (۱۹۹۸)، افعـال کامـل، افعـال کمکـی و قیـدها بـالاترین بسـامد عبـارات احتیاطآمیز را دارند. در این مقاله نتایج بر اساس رابطه ماهیت انواع روش تحقیق مـورد مقایسـه قرار گرفته است. در همین حوزهٔ مقالات زبان‌شناسی کاربردی، آثار متفاوتی بـه بررسـی نـوع و نقش عبارات احتیاطآمیز پرداخته‌اند، از جمله عطایی و صدر[۲](۲۰۰۸) بـا بررسـی بخـش بحـث و نتیجه‌گیری ۱۰۸ مقاله، بهنام، نعیمـی و درویـش‌زاده[۳](۲۰۱۲) بـا بررسـی بخـش نتیجـه‌گیـری ۱۰۰ مقاله و حسنی و فراهانی[٤](۲۰۱٦) با بررسی ٦۰ مقاله و مقایسـه جنسـیتی و برتـر دانسـتن مردان در کاربرد این عبارات، به بررسی نوع و نقش عبارات احتیاطآمیز پرداخته‌اند.

افشین و استاجی (۱۳۹۰) در پژوهش خود به بررسی و مقایسه بسامد عبـارات احتیـاطآمیـز در مقالات علمی و پژوهشی زبان فارسی در رشته‌های شـیمی و ادبیـات پرداختـه‌انـد. آن هـا از ۱۱٤ مقاله (۵۷ مقاله شیمی،۵۷ مقاله ادبیات) استفاده کردند که تعداد کل واژه‌ها در بخش بحـث مقالات شامل ۱۰۳۳۷۱ واژه است و در این تحقیق عبـارات احتیـاطآمیـز بـه شـش گـروه فعـل وجهی، فعل اصلی، قید، اسم، صفت و ارجاع تقسیم شده اسـت. نتـایج تحلیـل نشـان مـی‌دهـد توزیع عبارات احتیاطآمیز در دو رشته ادبیات و شیمی متفاوت می‌باشد. آن‌ها این تفـاوت را بـه دلیـل ماهیـت متفـاوت ایـن دو رشـته ذکـر کـرده‌انـد. همچنـین طباطبـایی‌لطفی، قاسـمی و سادات‌طباطبایی (۱۳۹۸) به بررسی چکیده‌های مقالات رشته‌های علوم‌انسـانی و پایـه/ پزشـکی پرداخته‌اند. آن‌ها از ٦۰ چکیدهٔ مقالات علمی- پژوهشی در بررسی خـود اسـتفاده کـرده‌انـد. بـر

---

۲۱ Hashemi & Shirzadi

۲۲ Atai & Sadr

۲۳ Behnam, Naeimi & Darvishzade

۲٤ Hasani & Farahani

اساس مدل ایوانز نتیجه می‌گیرند که به جز دو زیرمرحله، سایر موارد در نگارش چکیده بین گروه‌های مورد مقایسه یکسان هستند. در هر دو پژوهش اگر با استفاده از پیکره زبانی مقالات بیشتری را مورد بررسی قرار می‌دادند، قابلیت تعمیم نتایج آن‌ها بیشتر می‌شد.

علوی‌نیا (۱۳۹۰) به بررسی عبارات احتیاط‌آمیز و تشدیدکننده در مناظره‌های تلویزیونی ریاست جمهوری ایران و امریکا پرداخته‌اند، زیرا نامزدهای انتخاباتی جهت دستیابی به اهداف خویش و جلب‌نظر مخاطبان از تدابیر کلامی ویژه خود بهره می‌جویند و این تدابیر نقش بسزایی در گفتار سیاسی دارند.

شرفی (۱۳۹۰) به بررسی تأثیر جنسیت به عنوان عاملی اجتماعی بر کاربرد عبارات احتیاط-آمیز در گفتار مردان و زنان فارسی‌زبان پرداخته‌اند. جامعه آماری پژوهش شامل ٤٨ دانشجوی کارشناسی ارشد دانشگاه بیرجند (٢٤ مرد و ٢٤ زن) در سال تحصیلی ۹۰ – ۸۹ بوده است. نتایج نشان داد تفاوت آشکاری از نظر بسامد عبارات احتیاط‌آمیز بین مردان و زنان وجود ندارد و این تفاوت تنها از نظر کارکرد این عبارات مطرح است. دوگانی (۱۳۹۱) بیان می‌کنند عباراتی همچون «شاید»، «اگر» یا «ممکن است» از مواردی هستند که باعث سلب اطمینان از گزاره مطرح شده می‌شوند. او در ششصد آیه از قرآن کریم به بررسی این نشانه‌ها بر اساس چارچوب هایلند (۲۰۰۵) پرداخت. به این منظور هفت سوره از قرآن را به صورت تصادفی انتخاب کرد که مشتمل بر سیصد آیه مکی و سیصد آیه مدنی می‌باشند. نتایج این تحقیق نشان می‌دهد که وجود این عبارات احتیاط‌آمیز، منتسب به خداوند و اصول دین نبوده است، بلکه نسبت به بندگان آورده شده است و بازهم در چنین شرایطی نوعی تهدید و هشدار بوده و این عبارات احتیاط‌آمیز نوعی معنای تاکیدی و تنذیری پنهان را در خود داشته‌اند.

ممبینی (۱۳۹۲) نیز در پژوهش خود با استفاده از مدل تعامل هایلند (۲۰۰۵) به بررسی عبارات موضع‌نما (عبارات احتیاط‌آمیز) در نقد ۱۲۰ کتاب و مجله در زبان انگلیسی پرداخته-است. وی بیان کرده‌است که یکی از عوامل مهم شکل‌دهی بحث نقد کتاب، استفاده از عبارات موضع‌نما توسط نویسندگان (منتقدین) می‌باشد و این عبارات ابزاری برای ابراز احساسات شخصی، نگرش، داوری و ارزیابی نویسنده یا گوینده است. نتایج برگرفته از تحقیق حاکی از آن است که شیوه برگزیده توسط نویسندگان در به‌کارگیری عبارات موضع‌نما در نقد کتب بر حسب دیدگاه شناختی و قوانین اجتماعی هر رشته‌ای متفاوت است.

بر اساس این پیشینه مفهوم عبارات احتیاطآمیز در زبان فارسی بر اساس چارچوب‌های مختلف و با در نظر گرفتن متغیرهای متفاوتی مورد بررسی قرار گرفته است. اما این تحقیقات در سطحی محدود و غیر قابل تعمیم انجام شده‌اند، بنابراین کاربرد عبارات احتیاطآمیز با استفاده از پیکره گستره زبانی لازم به‌نظر می‌رسد.

۳-روش‌شناسی

هدف این پژوهش بررسی عبارات احتیاطآمیز در مقالات علمی-پژوهشی اعضای هیات علمی دانشگاه به روشی ساده و نظام‌مند و با استفاده از پیکره الکترونیکی محقق‌ساخته بوده است. در ادامه روش ساخت آن همراه با روش تحلیل داده‌ها ذکر شده است.

۳-۱ روش جمع‌آوری داده‌ها

برای نیل به اهداف این تحقیق ابتدا نویسندگان همراه با تیم تحقیقاتی خود پیکره متون علمی اعضای هیات علمی دانشگاه فردوسی مشهد مشتمل بر ۱۱۰۰ مقاله و ۷۱۵٤۲۰۲ واژه را بوجود آوردند. پیکره اصلی شامل تمام مقالات چاپ شده اعضای هیات‌علمی دانشگاه فردوسی مشهد از ۱۳۸۸تا پایان سال ۱۳۹۷ در مجلات فارسی می‌شود. پیکره متون علمی اعضای هیات‌علمی دانشگاه فردوسی مشهد با عنوان " پیکره زبانی آکادمیک فردوسی" به مقالات اعضای هیات‌علمی دانشگاه فردوسی مشهد در حوزه‌های علوم‌انسانی، مهندسی، علوم-پایه و علوم‌کشاورزی پرداخته و شامل گروه‌های آموزشی زیر می‌شود: زبان و ادبیات فارسی، زبان و ادبیات عربی، زبان فرانسه، زبان انگلیسی، زبان روسی، زبان‌شناسی، علوم اجتماعی، جغرافیا، تاریخ، مهندسی کامپیوتر، مهندسی برق، مهندسی مکانیک، مهندسی عمران، مهندسی شیمی، مهندسی متالورژی و مواد، مهندسی صنایع، اقتصاد کشاورزی، بیوتکنولوژی وبه نژادی گیاهی، مکانیک بیوسیستم، علوم خاک، اگروتکنولوژی، علوم باغبانی و مهندسی فضای سبز، علوم دامی، علوم و صنایع غذایی، گیاه‌پزشکی، علوم مهندسی آب، فلسفه و حکمت اسلامی، فقه و مبانی حقوق اسلامی، علوم قرآن و حدیث، تاریخ و تمدن اسلامی، ادیان و عرفان تطبیقی، معارف اسلامی، حقوق جزا و جرم شناسی، حقوق خصوصی، علوم سیاسی، علوم پایه، پاتوبیولوژی، بهداشت مواد غذایی و آبزیان، علوم درمانگاهی، بهداشت و پیشگیری از بیماری های دامی، فیزیک، شیمی، زمین‌شناسی، زیست‌شناسی، اقتصاد، مدیریت، حسابداری،

روان‌شناسی، روان‌شناسی مشاوره و تربیتی، علم اطلاعات و دانش شناسی، مبانی تعلیم و تربیت، مدیریت آموزشی و توسعه منابع انسانی، مطالعات برنامه درسی و آموزش، آمار، رفتار حرکتی و مدیریت ورزشی، فیزیولوژی ورزشی و آسیب‌شناسی، ورزش و حرکات اصلاحی، تربیت بدنی و علوم ورزشی، معماری، شهرسازی، محیط زیست، مرتع و آبخیزداری، مدیریت مناطق خشک و بیابانی، علوم و مهندسی شیلات، و مهندسی طبیعت. پیکره زبانی آکادمیک فردوسی از طریق لینک مقابل قابل دسترسی است: (https://github.com/Text-Mining/Ferdowsi-Annotated-Academic-Linguistic-Corpus). بعد از گردآوری مقالات، با استفاده از نرم‌افزار پیکره‌ساز طراحی شده، داده‌ها به صورت خودکار از نظر مقوله-های گفتار از جمله اسم، فعل، صفت، حرف اضافه و قید برچسب‌گذاری شدند. در نهایت بر روی خروجی اصلی برچسب خورده، امکان جستجوی سریع واژگان و اصطلاحات ایجاد شد. جستجوها در پژوهش حاضر بر اساس رشته و دانشکده به دو دسته کلی حوزه علوم‌انسانی و حوزه‌های علوم‌پایه تقسیم شدند. تعداد واژگان این پیکره (حوزه علوم‌انسانی و علوم‌پایه) شامل ٤٧٥٣٢٢٩ واژه و تعداد مقالات بررسی شده، ٧٩٧ می باشد. لازم به ذکر است که دقت برچسب گذاری پیکره ٩٦٪ است. این پیکره قادر به برچسب گذاری مقولات، ریشه یابی کلمات، و تجزیه آن‌ها می‌باشد.

۲-۳ ابزار تحلیل

متاسفانه ابزارهای استاندارد پیش‌پردازش ایجاد شده برای متون زبان فارسی از قبیل شمس‌فرد[25] (۲۰۱۰)، سرابی، مهیار و فرهودی (۲۰۱۳) و سراجی، مقیسی[26] و نیور[27] (۲۰۱۲) به‌صورت رایگان منتشر نشده‌اند. برخی از ابزارهای کدباز[28] پیش‌پردازش موجود از قبیل خلش و ایمانی[29] (٢٠١٤)، جلیل‌نژاد، محمودی و دهداری[30] (۲۰۱۰) و منشادی (۲۰۱۵) نیز از دقت

---

۲۵ Shamsfard

۲۶ Megyesi

۲۷ Nivre

۲۸ Open Source

۲۹ Khallash & Imani

۳۰ Jadidinejad, Mahmoudi & Dehdari

مناسبی برخوردار نیستند. در نتیجه ابزار مورد نیاز در ایـن پـژوهش نـرمافـزار سـاخت پیکـره و نرم‌افزارهای تحلیل آن است که نویسندگان مقاله آن را طراحی کرده‌اند.

در ابتدا همهٔ نویسه‌های (حـروف) مـتن بـا معـادل اسـتاندارد آن جـایگزین و یکسان‌سـازی گردید. سپس ریشه‌یابی واژگان انجام شـد. هـدف الگوریتم‌هـای ریشـه‌یابی، حـذف پیشـوند و پسوندهای کلمات و تعیین ریشه اصلی کلمه هستند. همچنین برای بررسـی آمـار تعـداد تکـرار واژگان و برچسب‌های مختلف بصـورت مجـزا یـا در کنـار هـم و اسـتخراج قـوانین نحـوی و ساخت‌واژی در متون زبان فارسی از پیکره استاندارد دادگان[۳] استفاده شد.

ابتدا نرم‌افزار پیکره‌ساز آماده و بر اساس آن کل مقالات به شکل پیکره درآمدند و متـون بـه اجزای سازنده‌شان مانند جمله و ریشه کلمات تجزیه شدند. همچنین مقوله‌های واژگـانیماننـد صفت، فعل، حرف اضافه، قید و اسم نیز مشـخص شـدند. در نهایـت نـرم‌افـزار کمکـی بـرای استخراج اطلاعات طراحی شد و همچنین کلمات مرتبط با عبارات احتیـاط‌آمیز کـه بـر اسـاس معیار سلاجرمیر مشخص شده بودند، برای نرم‌افزار تعریف گردید و با اسـتفاده از آن نـرم‌افـزار استخراج شد. گرچه قابلیت به‌کارگیری پیکره بدست آمده بسیار بالا است، اما در ایـن پـژوهش صرفا برای مقایسه عبارات احتیاط‌آمیز در بین دو پیکره اسـتفاده شـده اسـت تـا کـارآیی پیکـره طراحی‌شده سنجیده شود. در مراحل بعد این پیکره برای انجام سایر تحقیقـات در اختیـار سـایر پژوهشگران قرار خواهد گرفت.

٤-تحلیل داده‌ها

همان‌گونه که در مقدمه ذکر گردید برای انجـام ایـن تحقیـق از دسـته بنـدی سـلاجرمیر از عبارات احتیاط‌آمیز استفاده شده است. نمونه‌ای از هر دسته در جدول ۱ ذکر شده است.

جدول ۱- مثال از کاربرد عبارات احتیاط‌آمیز در داده‌ها

نوع عبارت احتیاط‌آمیز	حوزه علوم‌پایه	حوزه علوم‌انسانی
افعال کمکی	**می‌توان** گفت هرچه شرایط اقتصادی...	بر اساس این جدول **می‌توان** بیان نمود..

۳ Dadegan Treebank

شک و ارزیابی	**ممکن است** این پرسش مطرح شود که..	بروز نتایج متناقض **محتمل است**،
عبارات اسمی، قیدی و صفتی	حداکثر مقدار **ممکن** خود را پیدا می‌کند،	برای تولید حداکثر عملکرد **ممکن** با تولید مطلوب...
تقریب‌ها	**به‌سختی** قابل مقایسه می‌باشند.	**معمولا** همراه با قرایح واضحی بوده...
عبارات آغازین	–	این **حاکی از نگرش ما** نسبت به...
جملات شرطی	**اگر** علف‌کش‌های هورمونی بعد از ..	**اگر** به مقدار کافی مصرف شوند،
عبارات احتیاط‌آمیز مرکب	–	**منطقی به‌نظر می‌رسد که** رکورد...

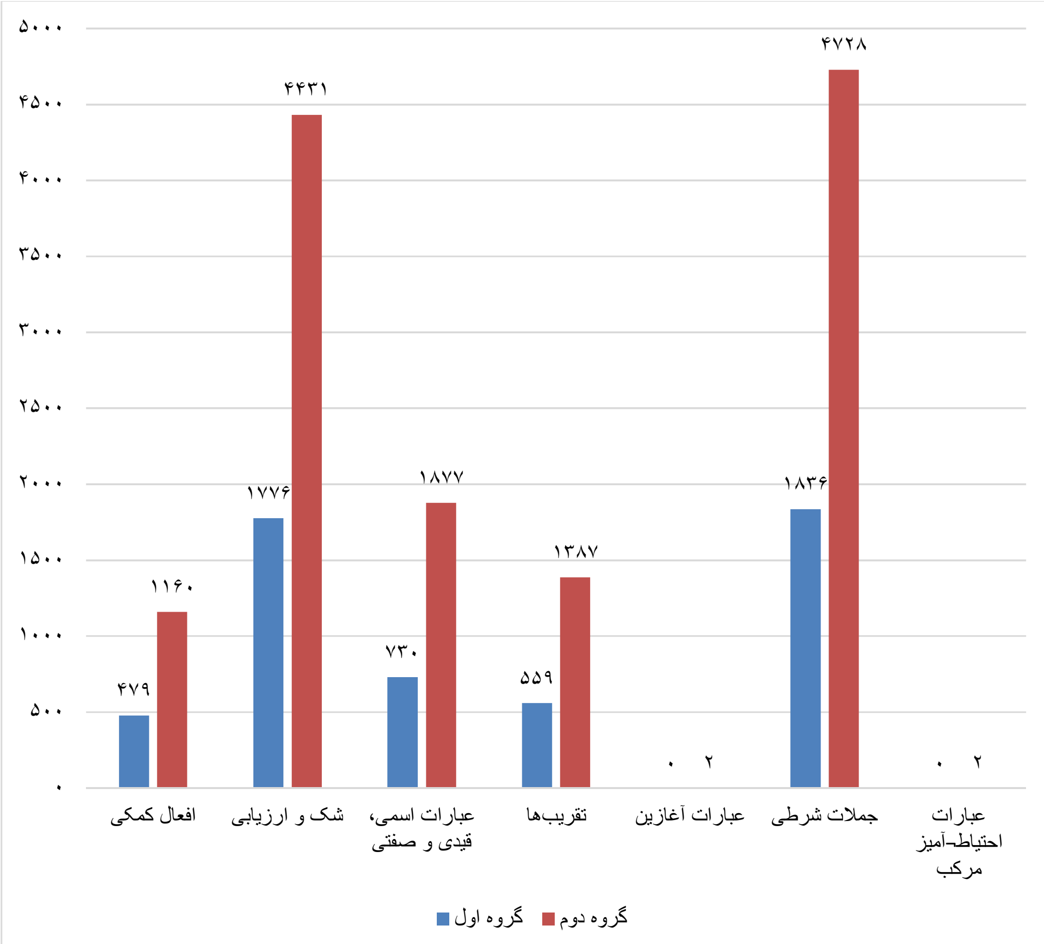
کل داده‌ها مورد بررسی قرار گرفتند و با استفاده از نـرم‌افـزار طراحـی شـده اطلاعـات هـر بخش جستجو شد و نتایج زیر (جدول ۲) به‌دست آمد.

جدول ۲- نوع و میزان استفاده عبارات احتیاط‌آمیز در مقالات دو گروه علوم‌انسانی و پایه

	حوزه علوم‌پایه (تعداد و درصد)	**حوزه علوم‌انسانی (تعداد و درصد)**
افعال کمکی	٤٧٩ (٥٪/٩)	١١٦٠ (٥٪/٨)
شک و ارزیابی	١٧٧٦ (٣٣٪)	٤٤٣١ (٣٢٪/٦)
عبارات اسمی، قیدی و صفتی	٧٣٠ (١٣٪/٥)	١٨٧٧ (١٣٪/٨)
تقریب‌ها	٥٥٩ (٪ ١٠/٣)	١٣٨٧ (٪ ١٠/٢)
عبارات آغازین	٠	٢ (٠٪)
جملات شرطی	١٨٣٦ (٣٤٪/١)	٤٧٢٨ (٣٤٪/٧)
عبارات احتیاط‌آمیز مرکب	٠	(٪ ٠/٢)
کل	٥٣٨٠ (٢٨٪/٣ از مجموع)	١٣٥٨٧ (٧١٪/٦ از مجموع)
مجموع	١٨٩٦٩	

به‌وضوح قابل مشاهده است که میـزان کـاربرد عبـارات احتیـاط‌آمیز در حـوزه علوم‌انسـانی حدود دو برابر بیشتر از حوزه دیگر است. در حوزه علوم‌پایه جمـلات شـرطی بـا تعـداد ١٨٣٦

(۳٤/۱ ٪) و پس از آن عبارات حاوی *شک و ارزیابی* بـا تعـداد ۱۷۷٦ (۳۳ ٪) بـالاترین میـزان استفاده را دارند و دو دسته عبارات آغازین و عبارات احتیاط‌آمیز مرکب بکار بـرده نشـده‌انـد. از نظر میزان استفاده، این ترتیب در حوزه علوم‌انسانی نیز مشاهده می‌شود، هـر چنـد میـزان آن بـه مراتب بیشتر از حوزه اول و حدود دو و نیم برابر است، تعـداد ٤۷۲۸ (۳٤/۷ ٪) بـرای جمـلات شرطی و تعداد ٤٤۳۱ (۳۲/٦ ٪) برای جملات حاوی شک و ارزیابی. در این حـوزه همـه انـواع عبارات احتیاط‌آمیز مشاهده شد، هر چند تنها دو نمونه از عبارات آغازین و عبارات احتیـاط‌آمیز مرکب در کل پیکره وجود داشت. همان‌گونه که در جـدول ۲ ذکـر شـده اسـت، در وسـط ایـن پیوستار افعال کمکی، عبارات اسمی، قیدی و صفتی و تقریب‌ها قرار دارند. موارد ذکـر شـده بـه صورت دقیق‌تر در نمودار زیر مشخص هستند. اینکه جملات شرطی و عبارات حاوی شـک و ارزیابی بالاترین میزان کاربرد در هر کدام از حوزه‌ها را دارند، انعکاسی از میـزان اسـتفاده بـالای این دو گونه در زبان است. مشـروط دانسـتن جـوابِ شـرط و شـک و ارزیـابی کـردن ماهیـت پژوهش‌های علمی هستند. اما پایین‌تر بودن میزان آن‌ها در علوم‌پایه نشانی از مسـتدل‌تر بـودن و قطعی‌تر بودن نتایج در این حوزه می‌باشد.



شکل ۱- مقایسه عبارات احتیاط‌آمیز در دو حوزه علوم‌انسانی و علوم‌پایه

۵-خلاصه و نتیجه‌گیری

برای انجام این پژوهش در ابتدا پیکره‌ای از مقالات علمی‌پژوهشی اعضای هیات علمی دانشگاه فردوسی مشهد ساخته شد. به‌طور خلاصه می‌توان گفت این عملیات در ٦ مرحله انجام شده است. اول جمع‌آوری متن مقالات و سپس تبدیل متون پیکره خام اولیه به فرمت قابل خواندن برای ماشین. سپس کار گردآوری منابع دستور زبان برای کار بر روی پیکره آغازین انجام شد. در دو مرحله بعد انواع برچسب‌ها به پیکره اضافه شد. در مرحله آخر استخراج دانش صریح از روی پیکره امکان‌پذیر شد. می‌توان گفت از ساده‌ترین موارد کاربرد

این پیکره، بررسی دادهها با شمارش جملات، واژههـا و تـکواژهـا اسـت. همچنـین اسـتخراج کلیدواژهها و واژههای همایند نیز از این دست هستند. بهطور دقیق، حوزههای مورد هـدف ایـن پیکره، مطالعات زبان برای اهداف ویژه و دانشـگاهی اسـت. بـه عنـوان مثـال بـر روی واژگـان رشتههای دانشگاهی، همنشینی واژگان، زنجیرههای واژگانی و روابط دستوری است. هـمچنـین با استفاده از این پیکره میتوان فهرستهای مختلفی از واژگان عمومی یا تخصصـی گـروههـای آموزشی مختلف استخراج کرد. این فرایند تاییدکننده آثاری چـون کاکسـهد[۳۲](۲۰۰۰)، گـاردنر و دیویس[۳۳](۲۰۱۳) است. همچنین با استفاده از این پیکره میتوان شـیوه نگـارش متـون علمـی در بین رشتهها و نویسندگان مختلف مورد بررسی قرار داد. از جملـه آثـاری کـه تاییـد کننـده ایـن مطلب است هایلند و تسو[۳۴](۲۰۰۴) است. آنچه در این پژوهش مد نظـر قـرار گرفـت، عبـارات احتیاطآمیز بودند. از طرفی دیگر، این پیکره از قابلیت پـردازش فایلهـای متنـی بـا فرمتهـای RTF، docx، doc، txt و pdf برخوردار است کـه واکـاوی داده را بـرای کـاربر بسـیار سـاده کرده و نیاز به تبدیل فایل ندارد. علاوه بر این، تولید این پیکرههای زبـانی، راه را بـرای بسـیاری از پژوهشهای زبـانی در حـوزههـای مختلـف از جملـه زبـانشناسـی رایانشـی، زبـانشناسـی اجتماعی، تحلیل متن و غیره هموار میکند و موجب کاهش زمان و هزینه و بالا بـردن دقـت در پژوهشهای آتی میشود.

عبارات احتیاطآمیز با هدف افزایش یا کاهش میزان تعهد، ابهام، وضوح، تحکیم یـا تضـعیف همبستگی در کلام بهکار مـیرونـد. در پـژوهش حاضـر ایـن عبـارات بـر اسـاس دسـتهبنـدی سلاجرمیر (۱۹۹۷) و بین دو حوزه دانشـگاهی علـومانسـانی و علـومپایـه مـورد بررسـی قـرار گرفتند. برای انجام آن پیکرههای زبانی مقالات علمیپژوهشی دانشگاه فردوسـی مشـهد جمـع-آوری و ساخته شد (پیکره زبانی آکادمیک فردوسی) و نتایج استخراج شد. اسـتخراج اطلاعـات کاملا به صورت خودکار انجام گرفت و تنها برای ارزیابی عملکرد پیکره و کـارایی آن، مـواردی به صورت نمونه تصادفی مورد ارزیابی دستی قرار گرفت که صحت اطلاعـات اسـتخراج شـده توسط نرمافزار ۹۶ درصد بود.

---

۳۲ Coxhead

۳۳ Gardner & Davies

۳۴ Hyland & Tse

بر اساس نتایج به‌دست‌آمده کاربرد عبارات احتیاط‌آمیز در حوزه علوم‌انسانی به مراتب بیشتر از حوزه علوم‌پایه بود و این نتیجه مؤید برخی پژوهش‌های انجام شده همچون افشین و استاجی (۱۳۹۰) و طباطبایی‌لطفی، قاسمی و سادات‌طباطبایی (۱۳۹۸)؛ فلاحتی (۲۰۰٤)؛ عبدی (۲۰۰۲)؛ سودمند افشار، مرادی و همزوی (۲۰۱۳)؛ و صحراگرد و یزدان‌پناه (۲۰۱۷) می‌باشد. آن‌چه که این پژوهش را از تحقیقات گذشته متمایز می‌کند، استفاده از پیکره در پژوهش و استفاده از پیکره ساخته‌شده و نرم‌افزار استخراج اطلاعات است. با بررسی داده‌ها می‌توان این‌گونه نتیجه‌گیری کرد که نویسنده می‌تواند با استفاده صحیح از عبارات احتیاط‌آمیز، خود را از هر گونه انتقاد احتمالی که در آینده بر ادعای علمی وی وارد می‌شود، مصون نگه دارد؛ زیرا از قبل بر عدم اطمینان از ادعای خود اقرار نموده است. اما همان‌گونه که مشاهده شد، نویسندگان حوزه‌های مختلف علمی به میزان متفاوتی از این عبارات استفاده کرده‌اند که به نظر می‌رسد این موضوع به ماهیت رشته‌های فوق‌الذکر برمی‌گردد. به طور مثال زمانی که در رشته‌ای از علوم‌پایه با آزمایش‌های دقیق و مکرر، کریستال‌های مختلف مورد بررسی قرار می‌گیرند، با اطمینان بیشتری می‌توان نتایج آن را ذکر کرد. در مقابل، تحلیل‌هایی در علوم‌انسانی وجود دارند که بیشتر بر اساس تجربه و شم زبانی نویسنده به‌دست آمده‌اند و نویسندگان که آگاه به الزام رعایت اصول نوشتار علمی هستند، با اقرار به عدم اطمینان کامل به نتایج یا مشروط بودن آن، با بکار بردن عبارات احتیاط‌آمیز این اصول را رعایت می‌کنند.

در واقع با بکارگیری این‌گونه عبارات، نویسندگان از دانش ارائه شده در مقاله فاصله می‌گیرند؛ در عین حال که بافت علمی احتمال هر گونه رنگ و بوی فرهنگی و شخصی را در مقالات منتفی کرده‌است، با وجود عبارات احتیاط‌آمیز پیشاپیش احتمال وجود خطا نیز که آن هم ناشی از ماهیت پژوهش است، پذیرفته شده است. زمانی که چیزی به طور مستقیم مشاهده می‌شود عباراتی همچون «دریافته‌ایم که..» بکار می‌رود و زمانیکه نتایج با تعبیر و تفسیر شخصی ارائه می‌شوند، جمله‌بندی به اشکالی مانند «اما ممکن است..» خود نمایی می‌کنند. به نظر می‌رسد استفاده از علم آمار و نتایج مستدل در آزمایشگاه‌ها که در حوزه علوم‌پایه بیشتر مشاهده می‌شود، روشی برای دور ماندن از عبارات احتیاط‌آمیز است. امروزه مشاهده می‌شود که در حوزه علوم‌انسانی نیز از دانشجویان و پژوهش‌گران خواسته می‌شود از تحلیل‌های آماری استفاده کنند و پژوهش‌هایی که صرفا بر اساس شم زبانی و تعبیر و تفسیر نویسنده است مورد

استقبال قرار نمی‌گیرند. در این پـژوهش بـه تفـاوت میـزان کـاربرد عبـارات احتیـاط‌آمیز در دو حوزه آموزشی پرداخته شد. در ادامه این تحقیق و به عنوان یک موضـوع پـژوهش می‌تـوان بـه بررسی میزان کاربرد عبارات احتیاط‌آمیز در حوزه علوم انسانی در دو فاصـله‌ی زمـانی پرداخـت و تاثیر استفاده از علم آمار در افزایش یا کاهش و میزان به‌کارگیری این گونـه عبـارات سـنجیده شود.

همان‌گونه که در پیشینه ذکر شد، پژوهش‌های پیشین بـا تمرکـز بـر دو رشـته بـه مقایسـه عبارات احتیاط‌آمیز پرداخته بودند، اما در پژوهش حاضر دو حوزه مختلف که هـر کـدام شـامل چندین رشته هستند، مورد مقایسه قرار گرفتند و بر خـلاف پـژوهش‌هـای پیشـین کـه بـر روی تعداد مقالات و واژگـان معـدود صـورت مـی‌گرفـت، پـژوهش حاضـر بـر روی ۷۹۷ مقالـه و ٤٧٥٣٢٢٩ واژه انجام شد و در نتیجه امکان تعمیم‌پذیری نتایج ایـن پـژوهش بـیش‌تـر اسـت. از این گذشته پیکره ساخته شده در فرآیند این پژوهش بستر پژوهش‌های گسترده در سایر حـوزه-های تحقیقاتی را فراهم کرده است. به زودی لینک استفاده از این پیکره بر روی سایت دانشـگاه فردوسی مشهد قرار داده خواهد شد تا سایر پژوهشگران که قصد انجام پژوهش بـر روی متـون دانشگاهی را دارند بتوانند بدون دغدغه‌ی جمـع‌آوری یـا جسـتجوی دسـتی داده‌هـا، بـا صـرف زمانی اندک مراحل بنیادی پژوهش خود را انجام دهند. در ضمن این پیکره متن بـاز اسـت و در نتیجه به طور منظم به روز رسانی خواهد شد.

کتابنامه

۱- افشین، فهیمه. و استاجی، اعظم. (۱۳۹۰). بررسی کاربردشناختی عبـارات احتیـاط‌آمیـز در مقالات علمی پژوهشی فارسی. ***پژوهش‌های زبانی***. (۲)۲. ۱۷– ۳٦.

۲- دانشکار آراسته، پویان. (۱۳۸۱). ***نرم‌افزار تشـخیص فعـل در زبـان فارسـی***. (پایـان‌نامـه منتشرنشدهٔ کارشناسی ارشد). تهران: دانشگاه علامه طباطبایی.

۳- دوگانی، شادی. (۱۳۹۱). ***بررسی عبـارات صـوری احتیـاط در هفـت سـوره قـرآن*** *(در چارچوب هایلند)*. (پایان‌نامه منتشرنشدهٔ کارشناسی ارشـد زبانشناسـی). مشـهد: دانشـکده ادبیات و علوم‌انسانی دکتر علی شریعتی دانشگاه فردوسی مشهد.

٤- شرفی، زهرا. (۱۳۹۰). ***بررسی کاربردشناختی عبارات احتیاط‌آمیز در گفتار زنان و مردان فارسی‌زبان***. (پایان‌نامه منتشرنشدهٔ کارشناسی ارشد زبان‌شناسی). مشهد: دانشکده ادبیات و علوم‌انسانی دکتر علی شریعتی دانشگاه فردوسی مشهد.

٥- طباطبایی‌لطفی، سید عبدالمجید؛ قاسمی، طاهره. و ساداتطباطبایی، مهین. (۱۳۹۸). مقایسه مراحل در چکیده‌های مقالات رشته‌های علوم‌انسانی و پایه/ پزشکی طبق مدل دادلی‌ایوانز. ***زبان‌پژوهی دانشگاه الزهرا***، (۳۰) ۱۱. ۹- ۳۰.

٦- علوی‌نیا، سید مجید (۱۳۹۰). ***تحلیل کاربردی عبارات احتیاط‌آمیز و تشدیدکننده در مناظره‌های تلویزیونی ریاست جمهوری ایران و امریکا***. (پایان‌نامه منتشرنشدهٔ کارشناسی‌ارشد). اهواز: دانشکده ادبیات و علوم‌انسانی دانشگاه شهید چمران.

٧- فرخ، ماندانا. (۱۳۸۱). ***بررسی ساختمان افعال ساده و مرکب فارسی و تدوین روش‌های سرواژه‌سازی به کمک رایانه***، (پایان‌نامه منتشرنشدهٔ کارشناسی ارشد). تهران: دانشگاه آزاد اسلامی واحد تهران مرکزی.

٨- ممبینی، پگاه. (۱۳۹۲). ***بررسی عبارات موضع‌نما در نقد کتب مجلات: مقایسه میان-رشته‌ای زبان‌شناسی کاربردی و فیزیک***. (پایان‌نامه منتشرنشدهٔ کارشناسی ارشد). اهواز: دانشکده ادبیات و علوم‌انسانی دانشگاه شهید چمران.

٩- Abdi, R. (2002). Interpersonal metadiscourse: An indicator of interaction and identity. *Discourse Studies, 4*, 139–145.

١٠- Atai, M. & Sadr, L. (2008). A cross-cultural genre study on hedging devices in discussion section of applied linguistics research articles. *Teaching English language and literature*, *7*(2), 1-22.

١١- Behnam, B., Naeimi, A. & Darvishzade, A. (2012). A comparative genre analysis of hedging expressions in research articles: Is fuzziness forever wicked. *English Language and Literature Studies, 2*(2), 20-38.

١٢- Clemen, G. (1997). The concept of hedging: origins, approaches, and definitions. In R. Markkanen, & H. Schröder (Eds.), *Hedging and discourse: Approaches to the analysis of a pragmatic phenomenon in academic texts* (pp. ٢٣٥–٢٤٩). □□□□□□: □□□□□r Hildebrand.

١٣- Coxhead, A. (2000). A New Academic Word List. TESOL Quarterly, 34(2), ٢١٣–٢٣٨.

١٤- Dousti, M. & Eslami Rasekh, A. (2016). ELT students' gender differences in the use of hedges in interpersonal interactions: A mixed method approach applied. *Journal of Applied Linguistics and Language Research*, *3*(1), 217-2٣١.

۱۵– Falahati, R. (2004). A contrastive study of hedging in English and Farsi academic discourse [Unpublished Master's thesis], University of Victoria.

۱٦– Fraser, B. (2010). Pragmatic competence: The case of hedging. In G. Kaltenboeck, (Ed.), *New Approaches to Hedging (pp. 13-3٤).* □□□□□□□: Emerald Publishing.

۱۷– Gardner, D. & Davies, M. (2013). A New Academic Vocabulary List. Applied Linguistics.

۱۸– Hashemi, M. R. & Shirzadi, M. (2016).The use of hedging in discussion sections of applied linguistics research articles with varied research. *The Journal of Teaching Language Skills (JTLS),35*(1), 31-56.

۱۹– Hassani, M. T. & Dastjani Farahani, M. (2016). A discourse analysis of gender differences in the use of hedging devices in applied linguistics research articles. *English Language Teaching, 1(1), 59-73.*

۲۰– Hyland, K. (1995). The author in the text: Hedging scientific writing. *Hong Kong Papers in Linguistics and Teaching, 18*, 33-42.

۲۱– Hyland, K. (1996a). Writing without conviction? Hedging in scientific research articles. *Applied Linguistics*, *17*(4), 433- 453.

۲۲– Hyland, K. (1996b). Talking to the academy: Forms of hedges in science research articles. *Written Communication*, *13*(2), 251-28.

۲۳– Hyland, K. (1998a). *Hedging in scientific research articles*. John Benjamins: Amesterdam.

۲٤– Hyland, K. (1998b). Boosting, hedging and negotiation of academic knowledge. *Text, 18*(3), 349- 382.

۲۵– Hyland, K. (2000). Hedges, boosters and lexical invisibility: Noticing modifiers in academic writing texts. *Language Awareness, 9*(4), 179- ۱۹۷.

۲٦– Hyland, K. & Tse, P. M. (2004). Meta discourse in academic writing: A reappraisal. *Applied Linguistics*, *25*(2), 156-177.

۲۷– Jadidinejad, A. H., Mahmoudi, F. & Dehdari, J. (2010). Evaluation of PerStem: A simple and efficient stemming algorithm for Persian, in workshop of the cross-language evaluation forum for European languages. *Springer*, 98- ۱۰۱.

۲۸– Jalilifar, A. (2006). All the way through the hedges: A corpus analysis of hedges in research articles. Paper presented at *7th Iranian conference on linguistics,* Tehran, Iran.

۲۹– Jalilifar, A. (2007). Hedging as a pragmatic strategy: Variations across disciplines and cultures. *TELL, 1*(3), 43-69.

۳۰– Khallash, M. & Imany, M. (2014). Hazm: Python library for digesting Persian text. [cited 2015; Available from: https://github.com/sobhe/hazm.

۳۱– Kong, C. C. K. (2006). Linguistic resources as evaluators in English and Chinese research articles. *Multi lingua, 25*, 183–216.

۳۲– Manshadi, M. (2015). Farsi Verb Tokenizer. 2013. [cited 2015; Available from: https://github.com/mehdi-manshadi/Farsi-Verb-Tokenizer.

۳۳– Marti´n, P. & Burgess, S. (2004). The rhetorical management of academic criticism in research article abstracts. *Text, 24*, 171–195.

٣٤- Millan, E. L. (2008). Epistemic and approximative meaning revisited: The use of hedges, boosters, and approximators when writing research in different disciplines. In S. Burgess, P. Marti´n-Marti´n. (Eds.), *English as an Additional Language in Research Publication and Communication (*pp. 65–82*).* Bern, Germany: Peter Lang.

٣٥- Sahragard, R. & Yazdanpanahi, S. (2017). English engagement markers: A comparison of humanities and science journal articles. *Language Art, 2*(1), ١١١-١٣٠.

٣٦- Soodmand Afshar, H., Moradi, M. & Hamzavi, R. (2014). Frequency and type of hedging devices used in the research articles of humanities, basic sciences and agriculture. *Procedia - Social and Behavioral Sciences*, 36, 70-74.

٣٧- Sarabi, Z., Mahyar, H. & Farhoodi, M. (2013). ParsiPardaz: Persian language processing toolkit. *Computer and Knowledge Engineering (ICCKE), 3rd International Conference on* IEEE, Mashhad, Iran.

٣٨- Rezaie, M. & Takii, S. (2014). Hedging expressions in English and Persian MA and PhD theses: The case of Iranian learners. *Mediterranean Journal of Social Sciences, 5*(20), 2339- 2349.

٣٩- Samaie, M., Khosravian, F. & Boghayeri, M. (2014). The frequency and types of hedges in research article introductions by Persian and English native authors. *Science Direct,* 98, 1678 – 1٦٨٥. □□□: ١٠.١٠١٦/□.□□□□□□.٢٠١٤.٠٣.٥٩٣

٤٠- Salager Meyer, F. (1994). Hedges and textual communicative function in medical English. *English for Specific Purposes*, *13*(2): 149- 170.

٤١- Salager Meyer, F. (1997). I think that perhaps you should: A study of hedges in written scientific discourse. In T. Miller (ed.), *Functional approaches to written texts: Classroom applications* (pp. 127–1٤٣). □□□□□□□□□□: □□□□.

٤٢- Sedaghat, A., Biria, R. & Asadi A.Y. (2015). Cross cultural analysis of hedges in Persian and English editorial columns. *International Journal of Language Learning and Applied Linguistics World*. *8*(1), 37-50.

٤٣- Seraji, M., Megyesi, B. & Nivre, J. (2012). A basic language resource kit for Persian. *Eight International Conference on Language Resources and Evaluation (LREC 2012),* European Language Resources Association. *Istanbul, Turkey*.

٤٤- Shamsfard, M. (2011). Challenges and open problems, Persian text processing. *5th Language & Technology Conference (LTC),* Poland.

٤٥- Shamsfard, M., Jafari, H. S., & Ilbeygi, M. (2010). STeP-١: □ □□□ □□ fundamental tools for Persian text processing. *Eight International Conference on Language Resources and Evaluation (LREC 2012),* European Language Resources Association. *Istanbul, Turkey*.

٤٦- Varttala, T. (2001). Hedging in Scientifically Oriented Discourse. Exploring Variation According to Discipline and Intended Audience.

٤٧- Vold, T. E. (2006). Epistemic modality markers in research articles: A cross linguistic and cross-disciplinary study. *International Journal of Applied Linguistics, 16*(1), 61-87.

Hedge Analysis in Ferdowsi Annotated Academic Language Corpus: A Comparison of the Fields of Humanities and Science

Atiyeh Kamyabi Gol (Corresponding author)
Assistant Professor, Department of Persian Language and Literature and Department of Linguistics, Ferdowsi University of Mashhad, Iran

Elham Akhlaghi Baghujeri
PhD, Department of Linguistics, Faculty of Literature and Humanities, Ferdowsi University of Mashhad, Iran

Hanieh Habibi
Department of Translation Technology, Faculty of Translation and Interpreting, University of Geneva, Switzerland

Received: 28/01/2021 Accepted:11/05/2021

A corpus consists of natural continuous texts that are presented electronically and from which information about linguistic elements (both lexical and non-lexical) can be extracted; this can be done in the shortest possible time and with the highest degree of accuracy. The authors of this article have created Ferdowsi Annotated Academic Language Corpus and the purpose of this article is to introduce the corpus of faculty members of Ferdowsi University of Mashhad and use it in analyzing and describing hedges in both fields of humanities and science. Hedges indicate the certainty of the author through words such as "maybe", "possibly" and "definitely". Research in this area has mainly focused on manual analysis of limited number of words. In this study, Salager Meyer’s (1997) classification has been used to identify, and categorize hedges. The data were extracted from Ferdowsi Annotated Academic Language Corpus which contains 1100 Persian language articles. Based on the results, the use of hedges in the field of humanities was about twice as much as in the field of science. Meanwhile, the use of conditionals and expressions of doubt in both fields had a high frequency. The results from this study have a better chance at being generalized to the fields under study due to the greater number of entries. The accuracy of the corpus labeling is ٩٦٪. [illegible]

scientific texts.

Keywords: Corpus, Scientific-research articles, Persian language, faculty members, Hedges